Building an Ontology-based Blockchain Application for Education

Shi Ming Jasmine KWOK

(3035787669)

Supervisor: Dr. Adela Sau Mui LAU

APAI3799: Directed Studies in Applied AI

The University of Hong Kong

10th May 2024

Acknowledgement

I would like to express my deepest appreciation to my advisor, Dr. Adela Lau, who has been the supervisor of my project. Her data science expertise profoundly influenced my research methodology and inspired me to explore innovative approaches to a revolutionary educational system. Her guidance regarding each step of the process, from initial conception to final implementation, has been invaluable.

Her patience and commitment to excellence have greatly motivated me, and her constructive criticism has been crucial in refining my arguments and methodologies. The insights and skills I have gained under her guidance are invaluable and will undoubtedly shape my future academic and professional endeavours.

I am truly thankful for her guidance and very fortunate to have had such a knowledgeable and empathetic supervisor. Thank you, Dr. Lau, for your unwavering support and belief in my abilities.

Abstract

The evolving landscape of global education demands innovative solutions to address the challenges of curriculum rigidity and the disconnect between educational output and market needs. This study explores the potential of blockchain technology as a transformative tool for the educational sector. By integrating decentralized and centralized blockchain architectures, we propose a novel educational model that utilizes Ethereum-based Non-Fungible Tokens (NFTs) and AI-driven predictive analytics to enhance curriculum flexibility and personalization.

Our approach involves the development of a dual blockchain system where a decentralized platform facilitates the creation, transaction, and verification of educational content as NFTs, ensuring the security and authenticity of materials. Concurrently, a centralized blockchain supports dynamic curriculum adjustments and real-time educational tracking, enhanced by AI capabilities that predict learning outcomes based on student interaction data. This integrated system not only safeguards intellectual property but also fosters a responsive and adaptive educational environment. The findings suggest that blockchain technology can potentially revolutionize educational practices by providing a more personalized, secure, and efficient learning experience.

Table of Contents

1. Introduction

1.1 Background

The landscape of education and curriculum development is undergoing a significant transformation, driven by the rapid evolution of the job market and emerging technologies. Traditional curriculums are increasingly seen as static and inflexible, often failing to meet the dynamic needs of both students and employers. These educational frameworks typically focus on singular disciplinary approaches, lacking the interdisciplinary integration required for addressing complex real-world problems. Moreover, there is a growing disconnect between the skills taught in educational institutions and those demanded in modern workplaces, leading to a skills mismatch that hampers both individual career success and organizational growth (Bridgstock, 2009).

To address these challenges, a shift towards a more adaptive, responsive educational model is necessary. The current centralized systems of curriculum development and content delivery are unable to adequately respond to these rapidly changing demands. Siemens (2015) and Hsieh, Yang, and Lin (2019) both advocate for a decentralized approach to education that enhances the flexibility of learning content and integrates cross-disciplinary knowledge with real-time market feedback.

1.2 Proposed Solution

Our proposal involves creating a blockchain application based on ontology that is specifically designed for the education sector with the goal of revolutionizing the current educational models.

This application will have two blockchain architectures: a decentralized integration with the Ethereum network to facilitate the creation and transactions of learning contents in the form of NFTs, and a centralized system to support dynamic curriculums with an AI component that predicts learning outcomes, trained using an ontology that encompasses multiple disciplines.

Centralized Blockchain Application: This component will act as the backbone for a dynamic curriculum framework set up on a centralized server. It will enable the real-time updating of the educational history of students to reflect their latest learning progress. Employers can review potential candidates' skill set and abilities through this application.

AI for Predictive Learning Outcomes: Integrated with the centralized blockchain, an AI system will analyse student performance and interaction data to predict future learning outcomes. This predictive model will be informed by a diverse ontology that includes various academic and professional disciplines, ensuring that the insights generated are comprehensive and beneficial for tailoring individual learning paths.

Ethereum-based Application for Content Exchange: To facilitate the exchange and verification of educational content, we will implement an Ethereum-based platform where learning materials are tokenized as Non-Fungible Tokens (NFTs). This approach not only secures the intellectual property rights of content creators but also ensures the authenticity and traceability of educational content across the blockchain.

Dedicated NFT Marketplace for Education: A specialized marketplace will be developed for the trade and exchange of educational NFTs. This marketplace will serve as a hub for educators, learners, and institutions to access, purchase, and sell educational content securely and transparently. This system will support a new ecosystem of educational exchange, fostering greater accessibility and innovation in learning material development.

1.3 Research Problem

Our research will focus on the integration of centralized and decentralized blockchain technologies to create a scalable and efficient NFT marketplace for educational content. The key research questions include:

- **Integration of Blockchain Technologies:** How can we effectively integrate centralized and decentralized blockchain systems to construct a NFT marketplace tailored for educational content?
- **Predictive Learning Outcomes:** What methodologies and models can be developed to accurately predict students' learning outcomes based on their interaction with blockchain-stored educational content?

1.4 Research Scope

The scope of this research will be based on the Environmental, Social, and Governance (ESG) criteria to ensure that the educational materials not only follow ethical and sustainable standards, but also successfully address the pressing global concerns of the day. The implementation of an ESG-focused strategy will direct the ontology's development and influence the kind of educational materials made available inside the blockchain ecosystem. This will facilitate the advancement of socially responsible and sustainable development goals-oriented education (Sachs, 2015).

2. Literature Review

2.1 Blockchain Development

2.1.1 What is Blockchain and Current Development and Its Limitation?

Blockchain technology is essentially a decentralized digital ledger that records transactions across multiple computers in such a manner that the recorded transactions cannot be altered retroactively. This technology is the backbone of cryptocurrencies such as Bitcoin and Ethereum, and its applications extend across various sectors including finance, healthcare, and supply chain management (Nakamoto, 2008).

Recent advancements in blockchain technology have focused on improving scalability and privacy. Innovations like the Lightning Network aim to enhance transaction speeds for cryptocurrencies like Bitcoin, while zero-knowledge proofs offer new ways of providing privacy by allowing transaction validation without revealing any underlying data (Buterin, 2014; Poon & Dryja, 2016). Despite these advancements, challenges remain, including issues with scalability, high energy consumption, and difficulties in integrating blockchain systems with existing legal frameworks, which can hinder broader adoption (Zheng et al., 2017; Deetman, 2016).

2.1.2 Centralized vs. Decentralized Blockchain

Centralized blockchains are managed by a single entity or a group of related entities. This centralization can lead to more efficient processing of transactions and streamlined governance. However, this central control also introduces potential security vulnerabilities and central points of failure, which can be exploited by attackers (Vukolić, 2015). Decentralized blockchains, on the other hand, distribute control across all network participants, enhancing security and resistance to censorship. This decentralization, however, can result in lower transaction throughput and increased latency, presenting a significant trade-off in network performance (Vukolić, 2015).

2.1.3 Non-Fungible Tokens

Non-Fungible Tokens (NFTs) represent a significant development in blockchain technology. NFTs are unique blockchain-based assets that signify ownership or proof of authenticity of a specific item or piece of content, distinguishing them from cryptocurrencies, which are fungible and can be exchanged on a one-to-one basis. The emergence of NFTs has notably impacted the realms of digital art, gaming, and collectibles, offering new mechanisms for digital ownership, content monetization, and creators' rights management (Wang & Kogan, 2021; Ante, 2021). NFTs facilitate a direct connection between creators and consumers, but they also raise concerns regarding valuation, copyright issues, and market sustainability (Dowling, 2022).

In the education sector, NFTs offer innovative ways to recognize and reward academic achievements and intellectual property. Here are a few potential applications:

- **Digital Certificates and Credentials:** NFTs can be used to issue verifiable digital certificates for academic achievements, professional qualifications, or completion of certain courses. These blockchain-based credentials are tamper-proof, easily verifiable, and portable across institutions and geographic boundaries, potentially enhancing the transparency and credibility of educational qualifications (Grech & Camilleri, 2017).

- **Student Portfolios:** NFTs can represent ownership over a student's portfolio of work, including projects, essays, art, and other intellectual efforts. This can provide a new way for students to maintain a verifiable and secure record of their academic and creative works, which can be shared with potential employers or educational institutions (Sharples & Domingue, 2016).
- **Reward Systems:** Educational institutions can use NFTs as a part of a reward system to motivate students. For example, achievements in certain academic or extracurricular activities could be recognized with unique NFTs, which could have practical benefits like access to advanced courses or resources.
- **Funding and Scholarships:** NFTs can be utilized to facilitate crowdfunding for scholarships or educational projects. Supporters can purchase NFTs, with the proceeds funding student scholarships or school initiatives, creating a community-supported model of educational finance (Hawlitschek et al., 2018).

In our project, we plan to utilize Non-Fungible Tokens (NFTs) as a novel approach to distributing and managing learning content. NFTs, with their unique capability to establish verified ownership and authenticity, will be used to represent individual pieces of educational material, such as modules, lessons, or multimedia content. This method ensures that each piece of content is distinctly identified on the blockchain, providing clear ownership and usage rights to the learners. By adopting NFTs in this manner, we aim to create a decentralized and secure system where educational resources are not only easily traceable but also protected against unauthorized use and reproduction. Furthermore, this approach could potentially enable a marketplace for educational content, where students and educators can trade resources in a transparent and secure environment, enhancing accessibility and encouraging the creation of high-quality educational material.

2.2 Ontology and Education

2.2.1 Introduction to Ontology and Its Use in Education

An ontology, in the context of information science, is defined as a formal and explicit specification of a shared conceptualization (Gruber, 1993). It provides a structured framework that not only helps in organizing information but also enhances the interoperability between various systems by enabling them to share and analyze knowledge in a coherent and accessible manner. This capability is particularly crucial in domains like education where the accurate and efficient dissemination of knowledge is key.

In educational settings, ontologies are increasingly recognized for their potential to revolutionize the way educational content is structured, accessed, and delivered. They facilitate the creation of curricular models that are adaptive and can be personalized to meet diverse learner needs. For instance, ontologies can be used to map the relationships between different learning modules, prerequisites, learning outcomes, and educational standards, providing a comprehensive framework that can be leveraged to design more effective educational programs (Dicheva & Dichev, 2006).

Furthermore, ontologies play a pivotal role in enhancing e-learning platforms. By categorizing and linking various educational resources and topics, ontologies enable more efficient search and retrieval of learning materials, making it easier for learners to find relevant content (Gašević, Djurić, & Devedžić, 2007). This structured organization of content also supports more sophisticated functionalities, such as recommendation systems that suggest learning resources based on the learner's progress, preferences, and learning objectives (Devedžić, 2006).

The use of ontologies in education also extends to facilitating personalized learning experiences. They enable the mapping of educational content to specific learner profiles, which includes their prior knowledge, learning styles, and educational goals. This approach not only enhances learner engagement by providing content that is tailored to individual needs but also improves learning outcomes by aligning educational materials to the learner's specific context (Knight, Gašević, & Richards, 2005).

In our project, we utilize ontological structures to enhance the prediction of learning scores for individual learners, leveraging the organized and interconnected nature of educational content. By defining an ontology that categorizes various educational concepts and their relationships—such as course prerequisites, topic complexity, and learning objectives—we can develop a more nuanced model of learner understanding and progression.

2.2.2 How to Build an Ontology

Building an ontology involves several steps: defining classes in the domain and the relationships among them, creating instances of classes, and continuously refining the ontology based on new information. Tools like Protégé offer platforms for ontology creation and maintenance (Noy & McGuinness, 2001).

In our project, we develop a 2-level ontology which involves creating a straightforward hierarchical structure where there are primary classes with a set of subclasses beneath each. This method of ontology design is particularly suited for domains where concepts can be clearly and easily classified into distinct categories, each with a more specific subset of elements.

To construct a 2-level ontology, the initial step involves defining broad, top-level classes that encapsulate the principal concepts of the domain. In a context that focuses on Environmental, Social, and Governance (E, S, and G) aspects, the top-level classes would be "Environment" (E), "Social" (S), and "Governance" (G). Each of these classes serves to capture a fundamental aspect of their respective domains without significant overlap.

Following the establishment of these top-level classes, each is further divided into subclasses to represent more specific concepts linked to the parent class. For example, under the "Environment" (E) class, subclasses might include "Climate Change," "Natural Resource Management," and "Sustainable Practices." These subclasses offer a finer level of categorization and aid in structuring the ontology to mirror more precise distinctions within the broader environmental category.

Similarly, for the "Social" (S) class, subclasses could be defined such as "Community Engagement," "Labor Practices," and "Equal Opportunity." Each subclass delves deeper into specific social issues, enhancing the ontology's ability to organize and reflect detailed social aspects.

Lastly, under the "Governance" (G) class, subclasses such as "Corporate Ethics," "Regulatory Compliance," and "Stakeholder Engagement" could be established. These subclasses provide a detailed breakdown of governance-related topics, enabling the ontology to accurately represent the nuances of governance within the broader category.

This 2-level structure maintains simplicity by limiting the depth of the hierarchy, which facilitates easier management and understanding of the ontology. It is particularly effective in scenarios where the domain knowledge can be naturally grouped into clear, distinct categories without requiring extensive interconnections among them. Tools like Protégé can be used to visually organize these classes and subclasses, making it easier to edit and refine the ontology as new insights or requirements emerge.

2.3 Topic Classification

Topic classification is a fundamental task in natural language processing (NLP) where text documents are categorized into predefined topics or classes. Various machine learning algorithms can be employed for this purpose. Here, we discuss three popular models: Naive Bayes, Word2Vec with SVM, and BERT (Bidirectional Encoder Representations from Transformers).

Naive Bayes Classifier: This probabilistic classifier is based on applying Bayes' theorem with strong independence assumptions between the features. It is particularly known for its simplicity and efficiency in handling large datasets. However, its assumption that all features are independent can limit its effectiveness in more complex text classification tasks where contextual relationships are crucial (McCallum and Nigam, 1998).

Word2Vec with SVM (Support Vector Machine): Word2Vec provides a dense word embedding that captures semantic meanings of words, which can then be used as features for a classifier like SVM. This combination benefits from the semantic understanding from Word2Vec and the effective classification capabilities of SVM, especially in linearly separable data. However, this method can be computationally intensive and may not scale well with very large vocabularies (Mikolov et al., 2013).

BERT (Bidirectional Encoder Representations from Transformers): BERT represents a breakthrough in transfer learning in NLP. It uses a transformer architecture to process words in relation to all the other words in a sentence, rather than one-at-a-time in order. This allows BERT to capture deep contextual relationships within text. Its pre-trained model can be fine-tuned with just one additional output layer to create state-of-the-art models for a wide range of tasks, including topic classification (Devlin et al., 2018).

For our text-to-class classification model, we chose BERT Base Uncased as our foundation due to its robust handling of context and its superior performance in various NLP tasks. The uncased model, which does not differentiate between uppercase and lowercase, reduces the complexity of the text input and is generally sufficient for understanding the semantics of educational texts where capitalization is less critical.

2.4 Types of Scoring Methods

Scoring methods for calculating learning outcomes, particularly in contexts like academic certificate attainment, vary widely. Each method offers unique advantages and faces distinct challenges. Here, we explore several different scoring methods:

Cumulative Grade Point Average (GPA): GPA is a widely used metric that averages a student's grades across their courses, weighted by credit points. Pros include its widespread recognition and ease of understanding for stakeholders. Cons involve its inability to reflect the depth of knowledge or skills in specific areas, and it may not account for varying course difficulties adequately (Rojstaczer and Healy, 2012).

Competency-Based Assessment: This method evaluates learning outcomes based on a student's ability to demonstrate specific competencies or skills rather than on traditional grades. Pros include a greater emphasis on practical skills and tailored learning paths. Cons are that it can be subjective, depending on the assessment criteria and the evaluator, and it may be challenging to implement uniformly across different educational institutions (Levine, 2005).

Rubrics and Descriptive Feedback: Rubrics provide detailed criteria for assessment and are often used alongside qualitative feedback. Pros include clear expectations and detailed feedback for improvement. Cons are that creating effective rubrics can be time-consuming, and they require significant training for evaluators to use consistently (Stevens and Levi, 2013).

In this project, we adopt an innovative scoring method that involves accumulating proportions of learned skills, a method we refer to as "Proportional Accumulation". This approach is designed to quantify learning in a more granular manner by tracking the incremental acquisition of skills and knowledge. When the accumulated proportion reaches or exceeds 1, it indicates that the student has fully mastered the skill. This method allows us to capture the nuances of learning progression in a way that traditional scoring systems like GPA might overlook.

The Proportional Accumulation method aligns with the principles of mastery learning, which suggest that students should progress to more advanced work only after achieving a set level of competency (Bloom, 1968). This method also takes cues from formative assessment strategies, where ongoing assessments are used to monitor student learning and provide continuous feedback that can guide future learning pathways (Black & Wiliam, 1998).

Critically, this scoring approach addresses some of the limitations of traditional and competency-based assessments. Unlike GPA, which may not adequately reflect the depth of knowledge or the difficulty of different courses (Rojstaczer & Healy, 2012), Proportional Accumulation provides a detailed view of student capabilities in specific areas. It also offers a structured framework that can mitigate some of the subjectivity associated with competency-based assessments (Levine, 2005).

Moreover, by utilizing a systematic approach to accumulate evidence of learning, this method can be standardized across various educational settings, addressing the implementation challenges noted with rubrics and competency-based systems (Stevens & Levi, 2013; Levine, 2005). This allows educators to implement a consistent and scalable method for assessing and recording student progress.

The Proportional Accumulation method provides a balanced approach by combining the clarity and structured nature of rubrics with the flexibility and skill-specific focus of competency-based assessments. This method not only enhances the accuracy of measuring learning outcomes but also supports personalized learning and feedback, crucial for educational environments aiming to foster deep and meaningful learning.

3. Research Methodology

In this section, we explain the methods we employ to tackle the research problems. We divide the development of the NFT Marketplace into 3 main components: horizontal blockchain management, vertical blockchain management, and the prediction model.

3.1 Horizontal Blockchain Management

In this component, we develop an API to access a database of blockchains where each blockchain represents a student's learning history. It should be able to retrieve and query blockchain information and data, update the database, and use the blockchain data for further processing such as learning score predictions.

According to a tutorial from GeeksforGeeks (n.d.), we can create a simple blockchain web application using Python. The basic structure of a block includes the index, timestamp, proof, and previous hash. A class called Blockchain is defined to handle the operations of the blockchain, including methods such as create_block which adds new blocks to the chain, proof_of_work which implements a simple proof-of-work algorithm to calculate the proof, hash which hashes the data of the block, chain_valid which checks the integrity of the blockchain, and a constructor which initializes an empty list to store blockchain blocks and creates the genesis block with a predefined hash of "0". The program utilizes Python Flask to provide routes /mine_block, /get_chain, and /valid via localhost on port 5000. We use this as the foundation of our API development.

In our implementation design, besides the index, timestamp, previous_hash, nonce (or proof), we also include a data field in a block to store the course title and course description for the learning scores prediction, and the transaction hash for referencing and linking up the vertical blockchain. Since we are maintaining a collection of blockchains, as opposed to a single blockchain in the GeeksforGeeks tutorial example, we should keep track of and be able to identify each of the blockchains from each another. An id field is therefore defined in the structure of a blockchain which is used for storing the wallet address of the blockchain owner, which uniquely identifies the blockchains. For simplicity, we will store all the blockchains in a Python list, instead of using database management systems such as MySQL or MongoDB. If performance and scalability issues are of concern, we can implement and migrate the data in the list to a database management system recommended above. The learning score predictions will be handled by the AI class which will be discussed later in section 3.3. The Flask application provides routes to create new chains (/create_chain), mine new blocks (/mine_block), get chains (/get_chain), get predicted learning scores (/get_prediction), and check the integrity of the blockchain (/check_valid) via localhost port 5000. The domain model below captures the key attributes and methods in our implementation design of the horizontal blockchain management API.

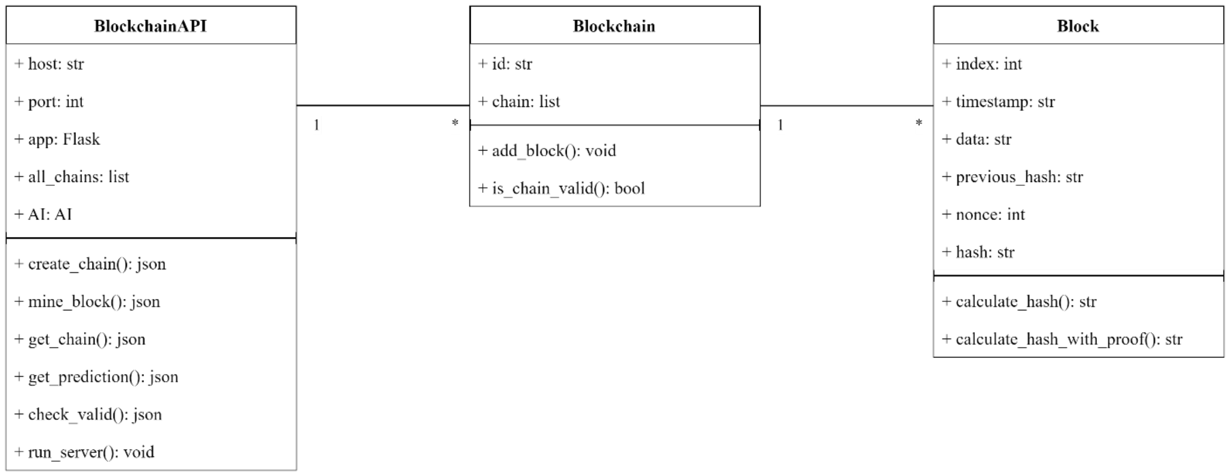

Figure 1. Domain model for Horizontal Blockchain Management

3.2 Vertical Blockchain Management

In this component, we develop an Ethereum-based smart contract as a framework to manage the transactions of learning contents in the form of NFTs. This allows for the creation, sales, and purchase of NFTs on the Ethereum network. We conceptualize this as the vertical blockchain management which orchestrates the transaction and ownership lifecycles of the NFTs on our marketplace.

Non-fungible Tokens, also known as NFTs, are a type of token that is unique in value and identity as introduced in Section 2.1.3. It gained popularity in fields such as digital art, video games, and music due to its non-interchangeable and tradable nature. We incorporate this architecture to our revolutionary education platform by introducing learning contents in the form of NFTs. Educators should be able to put up learning contents, specify the quantity, price, and royalty percentage in each NFT. For each sale, the creator of the NFT will receive a portion of the total sale price specified by the royalty percentage, while the seller will receive the remainder.

The smart contract we develop inherits from the ERC721 and ERC2981 standards, which are integral for managing NFT functionalities. ERC-721 is a standard for NFT which provides functionalities such as transferring tokens from one account to another, getting the current token balance of an account, getting the owner of a specific token, and the total supply of the token available on the network (Entriken, Shirley, Evans, & Sachs, 2018). Another standard for NFT used is ERC-2981, which provides a standardized way to retrieve royalty payment information to enable universal support for royalty payments across all NFT marketplaces and ecosystem participants (Burks, Morgan, Malone, & Seibel, 2020). We utilize these functionalities in our learning content trading system such that educators can create uniquely valued learning contents which supports royalty payments by deploying the smart contract on to the Ethereum network, and then learners can buy them from the platform or even re-selling them afterwards.

To achieve the vertical blockchain management, we incorporate the use of Etherscan as a third-party service to trace the transaction and ownership lifecycle of the NFTs. It provides a user-friendly interface to explore and search the Ethereum blockchain for transactions, addresses, tokens, prices, and other activities taking place on Ethereum. This enables us to retrieve details about a NFT collection including its original creator, total supply, and for each token minted, the platform can display a vertical representation of the transaction details such as the timestamp, buyer and seller addresses, and sales value, which are integral to achieve the vertical blockchain management.

In our development stage, we use Sepolia, a testnet of Ethereum, to simulate the deployment and transactions of the smart contract without spending real Ether with real market value. This testnet is crucial for developing and testing the functionalities of NFTs before deployment on the main network. To connect to the network, we require a Web3 provider such as Infura which provides us a connection node to the Sepolia network for smart contract transactions.

3.3 Prediction Model

In this component, we divide the workflow into 4 steps: building the ontology for ESG, preparing the dataset, fine tuning the BERT model for our topic detection task, and finally calculating the learning scores for ESG. Our main goal is to train a multi-label classifier which accepts textual data as inputs and outputs the topics detected from the texts, then apply our formula to calculate the learning scores of ESG.

Firstly, to build an ontology for ESG, we employ a simple methodology as suggested in Section 2.2.2 which starts by defining the surface level ontologies of the 3 domains: Environment, Social, and Governance, then proceed to define 10 subtopics for each of the domains. This creates a 2-level ontology which can offer basic classification of domain knowledge. To define the 10 subtopics for each domain, we use ChatGPT, a generative AI tool developed by OpenAI which is trained with a vast amount of data from diverse sources, to provide us with an unbiased and comprehensive insight. Our prompt to ChatGPT was “For each component in the ESG curriculum, list out 10 essential subtopics” and it replied:

ESG (Environmental, Social, and Governance) is a framework used to evaluate the sustainability and ethical impact of investments and companies. The curriculum for ESG typically covers various components related to these three pillars. Here are ten essential subtopics for each component:

Environmental:

1. Climate change and carbon footprint
2. Renewable energy and energy efficiency
3. Water management and conservation
4. Biodiversity and ecosystem preservation
5. Pollution control and waste management
6. Environmental regulations and compliance
7. Sustainable agriculture and land use
8. Green buildings and infrastructure
9. Environmental impact assessments
10. Supply chain sustainability and responsible sourcing

Social:

1. Human rights and labor practices
2. Diversity and inclusion in the workplace
3. Employee health and safety
4. Community engagement and development
5. Stakeholder management and accountability

6. Consumer protection and product safety

7. Social impact investing and philanthropy

8. Ethical marketing and advertising practices

9. Humanitarian initiatives and disaster response

10. Social equality and poverty alleviation

Governance:

1. Board structure and independence

2. Executive compensation and performance metrics

3. Shareholder rights and activism

4. Transparency and disclosure practices

5. Anti-corruption and bribery measures

6. Risk management and internal controls

7. Ethical decision-making and code of conduct

8. Political contributions and lobbying

9. Data privacy and cybersecurity

10. Stakeholder engagement and dialogue

These subtopics provide a broad overview of the key areas typically covered in an ESG curriculum. However, it's important to note that specific curricula may differ based on the educational institution or organization offering the program.

Secondly, to prepare the dataset, we need 30 columns to represent each of the ESG subtopics generated from ChatGPT as shown in the response in step 1. For each course, we label the respective columns as 1 to indicate the presence of the subtopics and 0 to indicate their absence. Another column stores the combined course information, including the title and description of the course. A total of 100 samples were gathered from various sources on the internet and were manually labelled to reflect the presence of the ESG subtopics based on the content described in the course titles and descriptions. We then split the dataset into 3 parts: 70% to the training set, and within the remain 30%, split them in half for the validation set and testing set. We randomize the datapoints to prevent order bias and add a random state to ensure reproducibility.

Thirdly, we adapt the “bert-base-uncased” model as the foundation of our prediction model. It is developed by Google as part of their BERT series, which stands for Bidirectional Encoder Representations from Transformers, which has been highly influential in the field of natural language processing (NLP) as mentioned in Section 2.3. We employ its pre-trained tokenizer for processing input data, ensuring that the text is appropriately formatted and encoded for BERT’s input requirements. To adapt the model for our topic detection task, an additional output layer consisting of a feed-forward neural network is integrated. This layer uses a sigmoid activation function to produce a prediction score between 0 and 1 for each label. This characteristic makes it particularly suitable for our multi-label classification task where each label is treated independently. Key hyperparameters include:

- MAX_LEN: 256,
- TRAIN_BATCH_SIZE: 32,
- VALID_BATCH_SIZE: 32,

- TEST_BATCH_SIZE: 32,
- EPOCHS: 10,
- LEARNING_RATE: 1e-05, and
- THRESHOLD: 0.5 (for sigmoid activation), are used to train our prediction model.

Finally, to calculate the ESG scores of a course based on the prediction results from our topic detection model, we first set some assumptions. Firstly, the prediction probabilities, which are the sigmoid output from the fine-tuned BERT model, implies the proportion of ESG subtopics taught in the courses. Secondly, even if courses overlap in content, repeated learning of subtopics implies a deeper understanding, thus accruing higher learning scores. Thirdly, a cap of 1 is set for the total scores for E, S, and G such that when the accumulated learning scores exceed 1, it indicates that the learner has already fully met the requirements of getting a certificate on E, S, or G. This aligns with our proposal on an innovative scoring system in Section 2.4 – “Proportional Accumulation”.

To integrate with the horizontal blockchain, we first extract the course titles and descriptions from the blockchain, then feed them as the input to the fine-tuned BERT model to detect relevant ESG subtopics within these courses. Normalize the prediction results to enhance the representation of the presence of subtopics, where weakly detected subtopics that have a prediction probability lower than the critical threshold will be removed. In other words, since our threshold for sigmoid activation is 0.5, each of the prediction probabilities will be subtracted by 0.5 and divided by 0.5 for normalization. Those that are less than 0 after normalization will be set to 0 to indicate no relevance to the subtopics. After that, sum up the probabilities for each domain E, S, and G, which is multiplied by a weight factor of 0.1, assuming that each subtopic is 1/10 of the total for each of E, S, and G, as each domain carries 10 subtopics. The sum is capped at a maximum of 1 for each domain to signify the completion of learning in that area.

4. Results

In this section, we report the results gathered from our research. This includes the overall software architecture for the NFT marketplace, and the evaluation on our fine-tuned BERT model for the topic detection task. See Appendix A for the source codes of the NFT marketplace, Appendix B for the source codes of the NFT smart contract, Appendix C for the source codes of the prediction model, and Appendix D for the dataset used in the model training and evaluation.

4.1 Software Architecture

The software architecture of our NFT marketplace integrates multiple components designed to handle blockchain management and data analysis through machine learning. The system's architecture is visualized through domain models and flowcharts that outline the interaction between these components as follows.

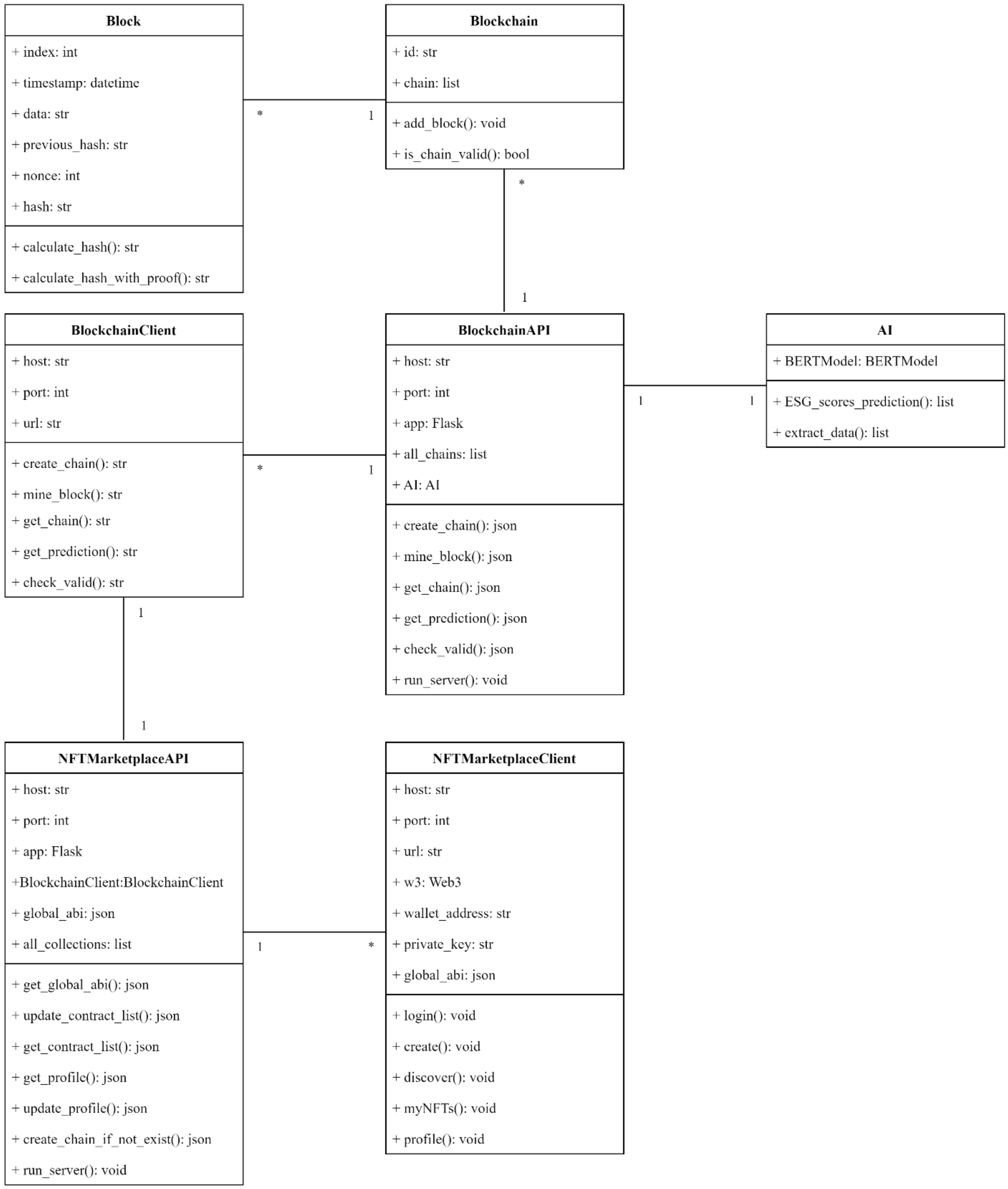

Figure 2. Domain Model for NFT Marketplace

Figure 2 shows the overall implementation design of the NFT Marketplace, which includes the horizontal blockchain management API mentioned in Section 3.1 with the addition of the AI class which handles the extraction of course details from the blockchain, get predictions from the fine-tuned BERT model, and calculation of the ESG scores based on the formula we discussed in Section 3.3. The BERTModel instance which is created in the AI class handles the loading of the pre-trained model, tokenization of the course titles and descriptions, and output the predictions for further processing.

The BlockchainClient is introduced to serve as an interface to send HTTP requests and get required information of the horizontal blockchain for the NFTMarketplaceAPI.

The NFTMarketplaceAPI will load the compiled smart contract in the format of ABI (Application Binary Interface) which is global to all NFT collections that are created on our marketplace, keep a list of deployed collections, and it is a Flask application that provides routes to update the contract list (/update_contract_list), get the list of contracts (/get_contract_list), update the learners' profile by adding block to their horizontal blockchain (/update_profile), get learners' profile (/get_profile), and create a new horizontal blockchain if a blockchain that corresponds to the wallet address does not exist on the blockchain server (/create_chain_if_not_exist) via localhost port 5500.

The NFTMarketplaceClient will verify the login wallet address and private key pairs, connect to the Ethereum network using the Web3 Python API for smart contract transactions, print a menu consisting of 4 options: Create NFTs, Discover NFT Collections, My NFTs, and My Profile, which handles the deployment of new smart contracts for NFT collections, minting NFTs, retrieving NFT collections and its tokens for purchase and listing, and a profile page which displays the horizontal blockchain showing the timestamps, course titles and descriptions, and the corresponding transaction hashes which can be copied and pasted to Etherscan for vertical blockchain management.

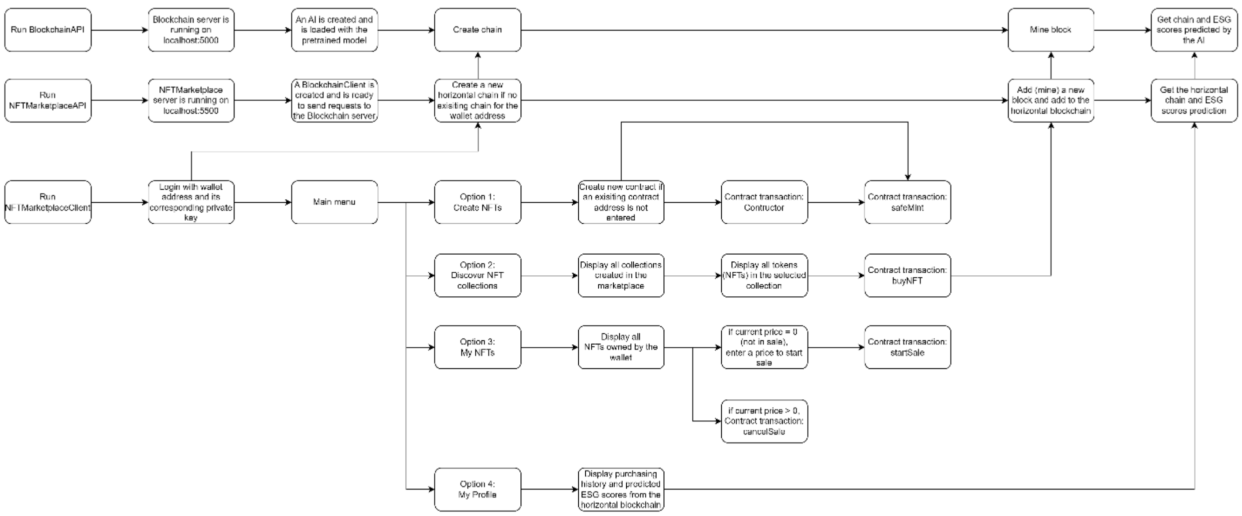

Figure 3. Flowchart for NFT Marketplace

Figure 3 shows the flowchart outlinling the step-by-step logical sequence of actions and decisions necessary for the execution of the program. BlockchainAPI and NFTMarketplaceAPI are required to execute before any clients are launched for the set up of the blockchain server, loading of the pre-trained BERT model, smart contract global ABI, and the list of contract addresses that corresponds to the NFT collections deployed on our marketplace.

Then, we can run NFTMarketplaceClient and login to a Ethereum wallet using a wallet address and private key pair which is validated using the Python eth_account library. A new horizontal blockchain representing the wallet address will be created if there is no record in the blockchain server. Upon successful login, the menu will be printed on command line and users can enter an option from 1 to 4.

If option 1 is chosen, the user will be prompted to enter an existing contract address that corresponds to a NFT collection that is created on our platform. Enter 0 to deploy a new contract to create a brand new NFT collection with specified name and symbol. The transaction hash of the deployment will be printed on screen and the user can now enter the details of the new NFTs, including the name, description, an optional external url for the course website, price, royalty percentage, number of NFTs, and an optional list of attributes. These information will be uploaded to the IPFS (InterPlanetary File System) in JSON format as defined in the ERC-721 standard for NFT metadata file. The transaction hash will be printed on screen after the miniting of new NFTs is confirmed. After the minting process is completed, the Python Web3 API will signal the program and the client goes back to the main menu.

If option 2 is chosen, the Web3 API will retrieve the NFT collection details from the Ethereum network by looking up the contract addresses stored in the list of contracts in NFTMarketplaceAPI one by one. The contract address, name, symbol, total supply, and the creator address will be printed on screen in tabular form. Then, the system prompts the user to enter a contract address to show the information of the tokens in a collection. Enter a contract address and the program will retrieve all the token information within a collection, including the token id, name, description, price, royalty information, and the current owner's address. These will be printed on screen in tabular form. Afterthat, the user can enter a token id to buy the NFT.

Upon successful transaction, the transaction hash will be printed on screen and a new block containing the course title, course description, and transaction hash will be added to the user's horizontal blockchain. The client goes back to the main menu when these operations are completed.

If option 3 is chosen, the program will fetch information of all the collections created in our marketplace to check if the current owner of any tokens matches the wallet address in the login session. Display all the token information for any matches found including the collection address, token id, name, description, price, and royalty information in which each of them are given an index starting from 1. Enter the index to start the sale of that NFT, specify a price and the transaction will proceed. If the NFT is already currently on sale, the program will cancel the sale so that other users can no longer buy it. The client goes back to the main menu afterwards.

If option 4 is chosen, the program will fetch the information of the wallet address's horizontal blockchain from the BlockchainAPI to get the learning history, predict the learning score using the fine-tuned BERT model, and also check the validity of the blockchain to ensure data integrity. The information will be displayed on screen in tabular format and the client goes back to the main menu.

4.1.1 Core Functionalities

The software was developed to support a range of functionalities integral to the management of blockchain data and the execution of a predictive model for educational content analysis. Key functionalities include:

- **Blockchain Interaction**: Capabilities to interact with both horizontal and vertical blockchains are fundamental. This includes creating and retrieving blocks, managing chain integrity, and executing transactions on the blockchain.
- **Smart Contract Deployment and Interaction**: The software automates the deployment of smart contracts on the Ethereum testnet and facilitates interactions with these contracts. This process is vital for managing NFT functionalities.
- **Data Extraction and Prediction**: The system extracts educational content data stored on the blockchain and employs a machine learning model to predict the relevance of this content to various ESG subtopics.
- **ESG Score Calculation**: After predictions, the system calculates an ESG score based on the presence of subtopics in the educational content, aiding stakeholders in assessing the content's comprehensive coverage of ESG criteria.

These functionalities were implemented using a combination of Python for backend operations and Solidity for smart contract development. Python's extensive library ecosystem, particularly for data science and blockchain, played a crucial role. Solidity was used for creating and managing smart contracts on Ethereum's blockchain.

4.1.2 Database and Data Handling

The database design is straightforward yet effective, tailored to the needs of blockchain data storage and manipulation. We utilized JSON files for simplicity and effectiveness in handling blockchain-related data:

- **Blockchain Data**: Blockchains and their constituent blocks are stored as JSON files. Each block and blockchain data structure includes relevant details such as block index, hash, and data, encapsulated in JSON format.
- **Smart Contract Data**: JSON files are also used to store the ABI (Application Binary Interface) and bytecode of smart contracts. This approach simplifies the deployment and interaction with contracts through our software.
- **Data Conversion**: Encoder and decoder scripts were developed to convert blockchain data between JSON strings and class objects. This conversion is crucial for ensuring that data manipulation within the software is efficient and that blockchain integrity is maintained.
- **IPFS Integration**: For handling NFT metadata, the IPFS (InterPlanetary File System) daemon is employed. This decentralized storage solution is ideal for managing the metadata files of NFTs, providing a robust and scalable storage solution.

4.2 Prediction Model Evaluation

The prediction model, which employs a BERT-based classifier, was evaluated to ensure its effectiveness in classifying educational content into various ESG subtopics.

4.2.1 Accuracy

The model was trained with a dataset of 100 labeled samples and tested with 15 samples. The overall test accuracy achieved was 74%, indicating a moderate to high level of precision in the model’s ability to correctly predict the presence of ESG subtopics in educational content, while keeping out of over-fitting.

The training and validation accuracies for each of the 10 epochs are captured in the line graph below. There is a upward trend throughout the training history which indicates improving model performance and good generalization.

Figure 4. Training history of the fine-tuned BERT model

4.2.2 Classification Metrics

In addition to accuracy, other classification metrics were computed to provide a comprehensive evaluation of the model's performance:

- **Precision**: The model showed an average precision of 0.51, suggesting that when it predicts the presence of a subtopic, it is correct 51% of the time.
- **Recall**: The average recall was 0.37, indicating that the model successfully identifies 37% of all relevant instances of each subtopic across the test data.
- **F1-Score**: The harmonic mean of precision and recall, the F1-score, was 0.42, reflecting the balance between precision and recall in our model.

Table 1. Classification metrics of the fine-tuned BERT model

	Precision	Recall	F1-score	Support
E1	0.50	0.33	0.40	3
E2	0.00	0.00	0.00	2
E3	0.00	0.00	0.00	3
E4	0.00	0.00	0.00	3
E5	0.00	0.00	0.00	3
E6	0.31	1.00	0.47	4
E7	0.00	0.00	0.00	2
E8	0.00	0.00	0.00	2
E9	0.00	0.00	0.00	4
E10	1.00	0.17	0.29	6
S1	0.46	1.00	0.63	6
S2	0.00	0.00	0.00	1
S3	0.00	0.00	0.00	6
S4	0.56	0.50	0.53	10
S5	0.78	0.70	0.74	10
S6	0.00	0.00	0.00	6
S7	0.69	1.00	0.82	9
S8	0.46	1.00	0.63	6
S9	0.00	0.00	0.00	3
S10	0.00	0.00	0.00	1
G1	0.00	0.00	0.00	2
G2	0.00	0.00	0.00	2
G3	0.00	0.00	0.00	6
G4	0.25	0.25	0.25	4
G5	0.00	0.00	0.00	1
G6	0.00	0.00	0.00	3
G7	1.00	1.00	1.00	1
G8	1.00	0.20	0.33	5
G9	0.00	0.00	0.00	0
G10	0.00	0.00	0.00	3
Micro average	0.51	0.36	0.42	117
Macro average	0.23	0.24	0.20	117
Weighted average	0.35	0.36	0.31	117
Samples average	0.51	0.37	0.42	117

We used the samples average approach to evaluate our model's performance, as we are dealing with datasets where each sample can belong to multiple classes simultaneously. This method involves computing the metrics for each individual sample and then averaging these metrics across all samples.

The results are poor with metrics lower than 50%. This suboptimal result may be attributed to the small size of the testing set, which comprises only 15 samples. Both precision and F-score have been set to 0.0 for labels where no predictions were made. Similarly, recall and F-score are also set to 0.0 for labels that do not have any true samples. This suggests issues with the model's ability to effectively predict certain classes, possibly due to the limited data in the testing set.

5. Discussion

In this section, we discuss the implications and outcomes of integrating blockchain technology within educational frameworks, as explored in this study. This includes the evaluation of the effectiveness and impacts of this integration, address the challenges faced, and the future potential of these technologies being applied in the education realm.

5.1 Evaluation of Blockchain and AI Integration in Education

The integration of blockchain technology in educational systems, as proposed in this study, offers a revolutionary approach to managing educational content and curriculum development. Central to our discussion is the dual architecture of blockchain— centralized horizontal blockchains and decentralized vertical blockchain—each serving distinct yet complementary functions.

The centralized blockchain component is instrumental in ensuring curriculum dynamism and real-time updates of educational progress. This system supports a more controlled environment that is necessary for sensitive operations such as updating student records and predicting learning outcomes based on AI analyses. Although the centralized nature may seem at odds with the traditional decentralized ethos of blockchain, it provides the necessary efficiency and swift response times that are crucial for academic administrative tasks (Swan, 2015).

On the other hand, the decentralized system, using the Ethereum network for the vertical creation and transactions of learning content as NFTs, emphasizes security, authenticity, and traceability of educational resources. This model taps into blockchain's inherent strengths in security and decentralized verification, effectively mitigating risks associated with content piracy and unauthorized access (Antonopoulos & Wood, 2018). Furthermore, by facilitating an NFT marketplace, it opens new economic avenues for content creators and institutions, potentially transforming the economics of educational content distribution (Tapscott & Tapscott, 2016).

The application of AI to predict learning outcomes is another cornerstone of this project. Integrated with the horizontal blockchains, the AI system analyses students’ study history to forecast their academic successes or skill levels. This predictive model, informed by an ontology that spans various academic and professional disciplines within our scope of ESG, has demonstrated potential in tailoring educational paths to individual needs more effectively than traditional methods (Russell & Norvig, 2016). This integration promises a more personalized and adaptive learning experience, aligning educational strategies with individual student profiles and learning trajectories (Hwang, 2014).

5.2 Challenges

While the integration of blockchain and AI holds promise, it has not been without challenges. The technological complexity and resource demands associated with maintaining blockchain systems are significant. Particularly, the energy consumption of blockchain operations, especially those based on Proof of Work (PoW) consensus mechanisms, poses environmental and economic challenges that need addressing to ensure sustainability (King & Nadal, 2012).

Furthermore, the use of a centralized blockchain architecture for certain administrative functions introduces security concerns. Since the system is hosted on HTTP servers, it becomes more vulnerable to attacks such as man-in-the-middle (MITM) attacks, where attackers can intercept and alter communications between two parties. This vulnerability highlights the need for enhanced security measures and potential migration to more secure protocols (Meiklejohn et al., 2013).

Moreover, our method of learning score prediction still requires further refinement to enhance its reliability and credibility. Our training results, as detailed in Section 4.2, show a test accuracy of 74%, precision of 51%, recall of 37%, and F1-score of 42%, indicating substantial room for improvement. The trade-offs made between dataset quality and quantity for our experiments underscore the challenges faced in balancing comprehensive software development with both the centralized and decentralized aspects of blockchain architecture.

Additionally, there is a lack of validation of our educational tools against real-world market and educational standards, which raises concerns about their credibility and applicability. Establishing relevance and efficacy in real educational settings is crucial to gain acceptance and widespread use. This lack of validation signifies a gap that must be bridged through rigorous testing and alignment with established educational benchmarks (Barabasi, 2016).

5.3 Future Directions

Moving forward, the project aims to refine the integration of blockchain technologies while expanding the AI predictive capabilities to include a broader range of data sets and learning environments. Continuous improvement of AI models is crucial to meet the evolving educational needs and adapt to changes in the dynamic job market, as noted by Zheng et al. (2017).

To address the environmental concerns associated with blockchain technology, future initiatives should focus on more energy-efficient consensus mechanisms, such as Proof of Stake (PoS). This change could significantly reduce the energy consumption of blockchain operations, thereby enhancing sustainability and reducing operational costs (King & Nadal, 2012). Additionally, enhancing the scalability and interoperability of blockchain systems is crucial for their broad adoption and long-term viability in educational sectors.

As we continue to develop and refine our blockchain and AI-enabled educational tools, addressing server vulnerability and scalability issues is paramount. The use of HTTP for hosting centralized blockchain components has exposed our systems to potential security threats, such as man-in-the-middle attacks. To mitigate these risks, future developments will focus on implementing more secure communication protocols, such as HTTPS, which encrypts data between the user and the server. Additionally, integrating advanced security measures like two-factor authentication and end-to-end encryption will further secure our systems against unauthorized access (Smith, 2018).

To improve the reliability and credibility of our learning score prediction methods, future development will focus on several key areas. Firstly, enhancing the quality and diversity of the datasets we use is essential. By incorporating a wider variety of data sources, including more granular educational outcomes and demographic information, we can train our models to better understand and predict learning behaviours across different populations (Jones et al., 2021).

Additionally, we plan to explore and integrate more advanced machine learning algorithms. Techniques such as deep learning and ensemble methods may provide significant improvements in model performance, increasing accuracy, precision, recall, and F1-scores beyond the current figures. These approaches can better capture complex patterns and interactions within the data, potentially leading to more robust learning predictions (Brown, 2022).

A critical step involves validating the Environmental, Social, and Governance (ESG) score formula by comparing the ESG scores generated by our system with those from established standard exams such as the GRE or SAT. This comparative analysis will ascertain the reliability and market relevance of our scoring mechanism, ensuring alignment with globally recognized benchmarks (Koretz, 2008).

6. Conclusion

This study embarked on an exploration of integrating blockchain technology within the educational sector, aiming to address the pressing issues of inflexibility and lack of interoperability in contemporary educational systems. By leveraging both centralized and decentralized blockchain architectures, we proposed a hybrid model that combines the robustness of AI-driven predictive analytics with the security and transparency of blockchain-based Non-Fungible Tokens (NFTs).

Throughout our research, we have demonstrated that the adoption of a blockchain-enabled educational model can significantly enhance the adaptability and responsiveness of curriculum development to meet the dynamic needs of both the job market and the educational demands of students. The decentralized aspect of blockchain allows for a flexible and real-time updating of educational content and credentials, which is essential in a rapidly evolving global job landscape.

Our implementation of an Ethereum-based platform for the creation and transaction of educational content as NFTs introduces a novel approach to managing and safeguarding intellectual property in the academic realm. This system not only ensures the traceability and authenticity of educational materials but also facilitates a marketplace for the exchange of these digital assets, thereby promoting a more dynamic educational ecosystem.

Moreover, the integration of AI to predict learning outcomes based on interaction with blockchain-stored content has paved the way for a more personalized and effective learning experience. This approach allows educational institutions to tailor educational paths that are aligned with individual student needs and market demands, thereby enhancing the overall educational impact.

However, challenges such as scalability, energy consumption, and legal implications of blockchain technology remain. Future research should focus on addressing these challenges to ensure broader adoption and to maximize the potential benefits of blockchain in education.

In conclusion, our research indicates that blockchain technology holds substantial promise for revolutionizing the educational sector by enhancing flexibility, security, and personalization. The potential for blockchain to support a decentralized, robust educational framework that aligns with future societal and technological shifts is significant, and it advocates for ongoing exploration and development in this promising field.

7. References

Ante, L. (2021). The non-fungible token (NFT) market and its relationship with Bitcoin and Ethereum. Frontiers in Blockchain, 4, Article 32.

Antonopoulos, A. M., & Wood, G. (2018). Mastering Ethereum: Building Smart Contracts and Dapps. O'Reilly Media.

Barabasi, A. L. (2016). Network Science. Cambridge University Press.

Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. Assessment in Education: Principles, Policy & Practice, 5(1), 7-74.

Bloom, B. S. (1968). Learning for Mastery. Evaluation Comment, 1(2), 1-12.

Bridgstock, R. (2009). The skills needed in the 21st century workplace: A career development perspective. Journal of Employment Counseling, 46(3), 118-127.

Brown, J. (2022). Advanced Machine Learning Techniques for Educational Data Analysis. Springer.

Burks, Z., Morgan, J., Malone, B., & Seibel, J. (2020). ERC-2981: NFT Royalty Standard. Ethereum Improvement Proposals. Retrieved from https://eips.ethereum.org/EIPS/eip-2981

Buterin, V. (2014). Ethereum: A next-generation smart contract and decentralized application platform. [White paper]. Retrieved from https://ethereum.org/en/whitepaper/

Deetman, S. (2016). Energy use for Bitcoin mining in emerging economies. Environmental Science and Technology, 50(24), 13577-13584.

Devedžić, V. (2006). Web Intelligence and Artificial Intelligence in Education. Educational Technology & Society, 9(4), 29-39.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv preprint arXiv:1810.04805.

Dicheva, D., & Dichev, C. (2006). TM4L: Creating and browsing educational topic maps. British Journal of Educational Technology, 37(3), 391-404.

Dowling, M. (2022). The cryptocurrency and NFT market: Developments, volatility, and financial stability implications. Journal of International Financial Markets, Institutions and Money, 75, Article 101437.

Entriken, W., Shirley, D., Evans, J., & Sachs, N. (2018). ERC-721: Non-Fungible Token Standard. Ethereum Improvement Proposals. Retrieved from https://eips.ethereum.org/EIPS/eip-721

Gašević, D., Djurić, D., & Devedžić, V. (2007). Ontology-based annotation of learning object content. Interactive Learning Environments, 15(1), 1-26.

GeeksforGeeks. (n.d.). Create simple Blockchain using Python. Retrieved May 8, 2024, from https://www.geeksforgeeks.org/create-simple-blockchain-using-python/

Grech, A., & Camilleri, A. F. (2017). Blockchain in education. In JRC Science for Policy Report. European Commission.

Gruber, T. R. (1993). A translation approach to portable ontology specifications. Knowledge Acquisition, 5(2), 199-220.

Hawlitschek, F., Notheisen, B., & Teubner, T. (2018). The limits of trust-free systems: A literature review on blockchain technology and trust in the sharing economy. Electronic Commerce Research and Applications, 29, 50-63.

Hsieh, Y. C., Yang, S. J. H., & Lin, C. S. (2019). Development and implementation of a decentralized education model: A case study. International Journal of Educational Technology in Higher Education, 16, Article 15.

Hwang, G. (2014). Designing and Implementing a Framework for Adaptive Learning Systems. International Journal of Advanced Computer Science and Applications, 5(3), 34-43.

Jones, S. M., Smith, P. L., & Roberts, J. (2021). Improving Educational Outcomes Through Data Diversity and Machine Learning. Journal of Educational Data Science, 3(2), 58-76.

King, S., & Nadal, S. (2012). PPCoin: Peer-to-Peer Crypto-Currency with Proof-of-Stake. Self-published white paper.

Knight, S., Gašević, D., & Richards, G. (2005). Ontologies to structure learning materials: An example of continuing professional development. Interactive Learning Environments, 13(1-2), 127-143.

Koretz, D. (2008). Measuring Up: What Educational Testing Really Tells Us. Harvard University Press.

Levine, A. (2005). Educating School Teachers. The Education Schools Project.

McCallum, A., & Nigam, K. (1998). A Comparison of Event Models for Naive Bayes Text Classification. AAAI-98 Workshop on Learning for Text Categorization.

Meiklejohn, S., et al. (2013). A Fistful of Bitcoins: Characterizing Payments Among Men with No Names. Communications of the ACM, 59(4), 86-93.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed Representations of Words and Phrases and their Compositionality. Advances in Neural Information Processing Systems, 26.

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. Retrieved from https://bitcoin.org/bitcoin.pdf

Noy, N. F., & McGuinness, D. L. (2001). Ontology Development 101: A Guide to Creating Your First Ontology. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880.

Poon, J., & Dryja, T. (2016). The Lightning Network: Scalable off-chain instant payments. [White paper]. Retrieved from https://lightning.network/lightning-network-paper.pdf

Rojstaczer, S., & Healy, C. (2012). Where A Is Ordinary: The Evolution of American College and University Grading, 1940–2009. Teachers College Record, 114(7), 1-23.

Russell, S. J., & Norvig, P. (2016). Artificial Intelligence: A Modern Approach. Pearson Education.

Sachs, J. (2015). The Age of Sustainable Development. Columbia University Press.

Sharples, M., & Domingue, J. (2016). The blockchain and kudos: A distributed system for educational record, reputation and reward. In European Conference on Technology Enhanced Learning (pp. 490-496). Springer, Cham.

Siemens, G. (2015). Connectivism: A learning theory for the digital age. International Journal of Instructional Technology and Distance Learning, 2(1), 3-10.

Smith, B. (2018). Securing Digital Transactions: Protocols and Practices. Wiley.

Stevens, D. D., & Levi, A. J. (2013). Introduction to Rubrics: An Assessment Tool to Save Grading Time, Convey Effective Feedback, and Promote Student Learning. Stylus Publishing, LLC.

Swan, M. (2015). Blockchain: Blueprint for a New Economy. O'Reilly Media.

Tapscott, D., & Tapscott, A. (2016). Blockchain Revolution: How the Technology Behind Bitcoin is Changing Money, Business, and the World. Penguin.

Vukolić, M. (2015). The quest for scalable blockchain fabric: Proof-of-work vs. BFT replication. In International Workshop on Open Problems in Network Security (pp. 112-125). Springer, Cham.

Wang, H., & Kogan, A. (2021). Designing NFT platforms to minimize agency costs. Journal of Emerging Technologies in Accounting, 18(2), 1-15.

Zheng, Z., Xie, S., Dai, H., Chen, X., & Wang, H. (2017). An overview of blockchain technology: Architecture, consensus, and future trends. In 2017 IEEE International Congress on Big Data (BigData Congress), (pp. 557-564). IEEE.

Zheng, Z., Xie, S., Dai, H., Chen, X., & Wang, H. (2017). An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends. In Proceedings of the 6th International Congress on Big Data, 557-564.

8. Appendices

GitHub repository: https://github.com/jkjk101/APAI3799

Demo video: https://github.com/jkjk101/APAI3799/blob/main/Demo.mp4

Appendix A. Source codes of NFT Marketplace

NFT Marketplace API: /NFTMarketplaceAPI.py

Blockchain API: /BlockchainAPI.py

NFT Marketplace Client: /NFTMarketplaceClient.py

Blockchain Client: /BlockchainClient.py

Blockchain: /Blockchain.py

AI: /AI.py

BERT: /BERT.py

Appendix B. Source codes of NFT Smart Contract

File path: /contracts/NFT.sol

Appendix B. Source codes of Prediction Model

File path: /models/multi_label_text_classification_BERT.ipynb

Appendix C. Dataset

File path: /models/courses.csv